BLIND DETECTION OF THE NUMBER OF COMMUNICATION SIGNALS UNDER SPATIALLY CORRELATED NOISE BY ICA AND K-S TESTS

Vincent Choqueuse

E^3I^2, EA 3876, ENSIETA
2, rue Francois verny, 29806 Brest Cedex 9
FRANCE

Koffi Yao, Ludovic Collin, Gilles Burel

LEST, UMR CNRS 6165, UBO
6, avenue le Gorgeu, 29200 Brest Cedex 3
FRANCE

ABSTRACT

The issue addressed in this paper is the determination of the number of communication signals in a sensor array. Most of the available algorithms rely on the spatial uncorrelation of the additive noise. In practice, this condition is rarely satisfied when the receivers are not sufficiently spaced (MIMO communications for example). In this paper, we propose a new method to detect the number of communication signals based on the fact that the signals are independent and non gaussian and that the background noise is gaussian. By using an Independent Component Analysis in conjunction with Kolomogorov-Smirnov (K-S) tests, the method can detect as many communication signals as the number of receiver antennas. Simulations results show that our method performs well in many environments like those with spatially correlated noise.

Index Terms— Signal Detection, MIMO Systems, Electronic Warfare

1. INTRODUCTION

The determination of the number of communication signals in a sensor array is a key-issue because of its applications in civilian and military telecommunication systems. Most of the algorithms available in the literature to fulfill this task are based on the assumption that the additive noise is spatially uncorrelated at the receiver side. The number of signals is then determined from the detection of the order of multiplicity of the lowest eigenvalues of the received-signals 2nd order statistic. To detect the order of multiplicity, some authors proposed to use hypothesis tests [1] or theoretical information criteria [2]. But the main drawback of all of these methods is that the detection is unsuccessful in spatially-correlated noise environments. In these kind of environment, some authors have suggested the use of two receiver-antenna networks [3]. On condition to have a large distance between both networks,

the noise between the two sensor arrays is assumed to be spatially uncorrelated. But this method implies the knowledge of the spatial correlation level to sufficiently move apart the two receiver-antenna networks. Furthermore, the number of receiver antennas must be high since these methods cannot detect a number of signals greater than half the number of receivers. To overcome the problem of 2nd order methods, Sawada has proposed a method based on a ICA [4] which works in correlated noise environments with frequency selective channels. The number of signals is obtained by comparing the normalized power of extracted components with a threshold. The main drawback of the method lies in the fact that the threshold is fixed beforehand when training data are available.

In this paper, we propose a new approach to blindly detect the number of communication signals in the case of a frequency flat channel. It relies on the assumption that the transmitted signals are independent and non gaussian and that the additive noise is gaussian. This report is organized as follows. Section II deals with the extraction of non-gaussian components (communication signals) from the received samples. This extraction step is realized by an Independent Component Analysis (ICA) algorithm. In section III, Kolmogorov-Smirnov (K-S) tests are used to detect the number of extracted components which can be considered as non-gaussian. Finally in section IV, the performances of our proposed method are compared to those of other algorithms.

2. EXTRACTION OF NON-GAUSSIAN COMPONENTS BY ICA

Let us consider a multi-transmitter communication system using n_t antennas. The communication is intercepted by a receiver composed of n_r antennas ($n_r \geq n_t$). Let us denote the n_r samples received at the time k by the column vector $R(k)$. Under the assumption that the channel is frequency flat and time invariant, the received samples can be expressed by:

$$R(k) = \mathbf{H}S(k) + B(k) \qquad (1)$$

where $\mathbf{H}$ corresponds to the $n_r \times n_t$ complex channel matrix and $B(k)$ is a $n_r \times 1$ column vector which corresponds to the additive complex gaussian noise (which may be spatially correlated). The $n_t \times 1$ column vector $S(k)$ corresponds to the transmitted complex symbols. The transmitted symbols are assumed to belong to a linear modulation (PAM, PSK, QAM) and to be spatially- and temporally-independent. The goal of this paper is to blindly detect the number of signals n_t from N observations of $R(k)$.

To detect the number of signals, an Independent Components Analysis (ICA) is performed on the received samples $R(k)$. Under the assumption that the number of signals n_t is unknown, an ICA algorithm finds a separating matrix of size $n_r \times n_r$, $\mathbf{W}$, which minimizes the gaussiannity of components, $Y(n) = [y_1(n) \cdots y_{n_r}(n)]^T$, so that:

$$Y(n) = \mathbf{W}R(n) \tag{2}$$

The reference [6] lists a large number of ICA methods proposed in the literature which can blindly find the matrix $\mathbf{W}$. In this paper we keep two algorithms, JADE [7] and Fast-ICA [8], which present good behaviors for signals far from gaussianity [9]. The JADE and ICA algorithms minimize the gaussianity of extracted components by maximizing respectively the 4th order cumulants and the negentropy of the components $Y(n)$. After convergence of one of these algorithms, the extracted components $Y(n)$ can be divided in two classes:

- Signal independent components. This class is composed of n_t components of unit variance, denoted $y_{t,l}(n)$ $(1 \leq l \leq n_t)$, which corresponds individually to a communication signals, $s_l(n)$, corrupted by unknown factors of scale and phase, α_l and θ_l respectively, and by a residual circular gaussian noise, $b_l(n)$.

$$y_{t,l}(n) = \alpha_l e^{j\theta_l} s_l(n) + b_l(n) \tag{3}$$

- Gaussian circular components. This class is composed of $n_g = n_r - n_t$ components for which the gaussiannity cannot be minimized. These components, denoted $y_{g,l}(n)$ $(1 \leq l \leq n_g)$, follow a gaussian circular law with a zero-mean and unit variance.

In the next section, we propose to detect the number of communication signals by determining the number of gaussian components $n_g = n_r - n_t$.

3. DETECTION OF NON-GAUSSIAN COMPONENTS WITH K-S TESTS

To detect the gaussianity of the independent component, we propose to use a Kolmogorov-Smirnov test which is based on

the cumulative distribution function (cdf).

By using the result of the previous section, the components $y_{g,l}(n)$ follow a gaussian circular law with zero mean and unit variance. Note that it is complex and computationally demanding to compare directly the 2D-cdf of the extracted components with the 2-D cdf of a gaussian circular law. By using the fact that the modulus of a zero-mean complex gaussian law with unit variance follows a Rayleigh law with parameter $b = \frac{1}{\sqrt{2}}$, we propose to restrict the comparison of $|y_i(n)|$ with a 1D Rayleigh distribution[1]. Let's consider the empirical distribution function, $F_i(x)$, of $|y_i(n)|$ $(i \in \{1, \cdots, n_r\})$ defined by:

$$F_i(x) = \frac{1}{N} \sum_{n=1}^{N} I_{|y_i(n)| \leq x} \tag{4}$$

where $I_{|y_i(n)| \leq x}$ is the indicator function. We compare $F_i(x)$ with the cumulative distribution function, $F(x)$, of a rayleigh law with parameter $b = \frac{1}{\sqrt{2}}$ defined by:

$$F(x) = \int_0^x 2te^{-t^2} dt \tag{5}$$

The figure 1 exposes the empirical cdf $F_i(x)$ for a MIMO communication using $n_t = 2$ emitters and $n_r = 3$ receivers. The Rayleigh cdf $F(x)$ is plotted on the same figure. We can remark that the cdf of $F_2(x)$ is the only one which fits well the cdf of the Rayleigh law and so $\widehat{n}_g = 1$. So, we can state that the number of signals is equal to $\widehat{n}_t = n_r - n_g = 2$.

To automatically detect the number of non-gaussian components we propose to perform a Kolmogorov Smirnov test on each component $y_i(n)$. The K-S test can be defined as follows:

- H_0: The data $|y(i)|$ follow a Rayleigh distribution of parameter $\frac{1}{\sqrt{2}}$, i.e. $F_i(x) = F(x)$.

- H_1: The data do not follow the specified distribution, i.e. $F_i(x) \neq F(x)$.

The K-S test is based on the maximum difference, D_i, between the empirical and the theoretical cdf.

$$D_i = \sup_x |F_i(x) - F(x)| \tag{6}$$

The hypothesis H_0 is rejected if $\sqrt{N}D_i$ is greater than a critical value T_α. For a continuous cdf F, the critical value is given from the significance level of the test α according to the following equation:

$$\alpha = 1 - Pr(\sqrt{N}D_i \leq T_\alpha) \tag{7}$$

---

[1]Note that our method does not exploit the phase information. Some improvements can be obtained by using the fact that the distribution of the phase for a gaussian circular variable follows an uniform law
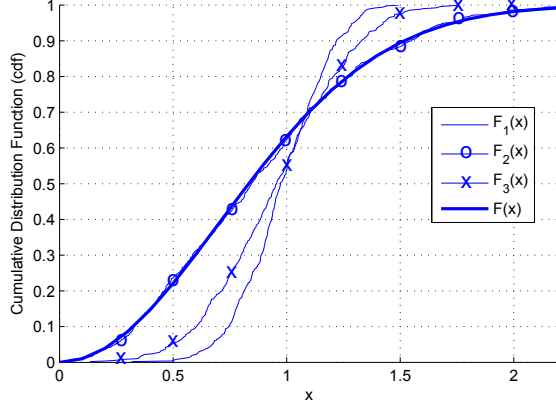
Fig. 1. Rayleigh-cdf and empirical cdf of the components $|y_i(n)|$ obtained with JADE-ICA algorithm. The communication uses 2 emitters which transmit 1000 symbols and 3 receivers. The communication signals are generated from a QPSK modulation and the SNR is equal to 5 dB.

When $N \rightarrow \infty$, $\sqrt{N}\mathrm{D}_i$ converges to the Kolmogorov distribution defined by:

$$Pr(\sqrt{N}\mathrm{D}_i \leq T_\alpha) = 1 - 2\sum_{k=1}^{\infty}(-1)^{i-1}e^{-2i^2T_\alpha^2} \quad (8)$$

Finally with a large number of received samples N, an estimation of the number communication signals $\widehat{n}_t$ is given by:

$$\widehat{n}_t = \sum_{i=1}^{n_r}\mathrm{U}(\sqrt{N}\mathrm{D}_i - T_\alpha) \quad (9)$$

where U is the Heaviside step function which is zero for negative argument and one for positive argument. Note that the proposed method can detect from 0 to n_r communication signals.

4. PERFORMANCE

To highlight the capability of our algorithm to detect the number of communication signals, we carried out several Monte Carlo simulations where the transmitted symbols always belonged to a QPSK modulation and the number of received samples was equal to 1000. At first, the algorithm under study was compared to the one proposed by Wax based on the use of Minimum Description Length (MDL) information criteria [2] in the case of spatially uncorrelated noise. Then, it was compared to the algorithms proposed by Chen [3] in the case of spatially correlated noise for a receiver composed of two arrays of antennas. For each simulation, we evaluate the performance of our proposed method with an ICA based on JADE and Fast-ICA.
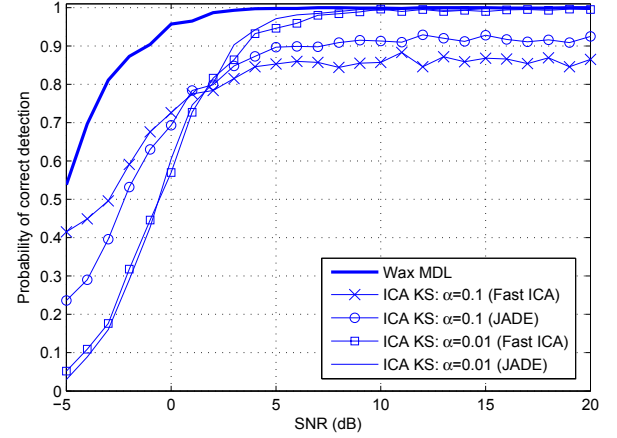

Fig. 2. Probability of correct detection of the number of sources($n_t = 2$) for a 2×2-antenna receiver and 1000 received samples per antenna; performances of the Wax and of the proposed algorithm.

4.1. Spatially uncorrelated noise environment

In the first tests, the additive noise is spatially uncorrelated, i.e. $E[B(k)B^H(k)] = \sigma^2\mathbf{I}$ where σ^2 is the noise variance. To assess the blind detection of communication signals by our algorithm, n_t was set equal to 2, and the receiver was a single array of 4 antennas ($n_r = 4$). The figure 2 compare the performances of our algorithm with different values α and the ones of the Wax algorithm. We remark that decreasing the value of α, decreases the performances of detection at low SNR but improves its for high SNR. We can note that our proposed algorithm based on Fast ICA is better than the one based on JADE ICA at low SNR but this trend is inverted at high SNR. Figure 2 also shows that, when the noise is spatially uncorrelated, the Wax algorithm performs better than our algorithm.

4.2. Spatially correlated noise environment

Different methods about the detection of a number of sources under conditions of spatially-correlated noise have been proposed in the literature. Among them, those developed by Chen and co-workers [3] use a receiver composed of two distinct Uniform Linear Arrays (ULA) of antennas. To compare the results produced by our algorithm against those of the algorithm reported in [3], each network was set to consist of 2 antennas, i.e. $n_r = 4$ and the spacing between the two array is twice the spacing between each antenna. Moreover, the spatially correlated noise is generated with a moving average filter of order 1 with a coefficient equals to 0.2. Figure 3 presents the probability of correct detection of the number of communications signals as a function of the SNR in the case of an emitter composed of $n_t = 2$ antennas. We remark that
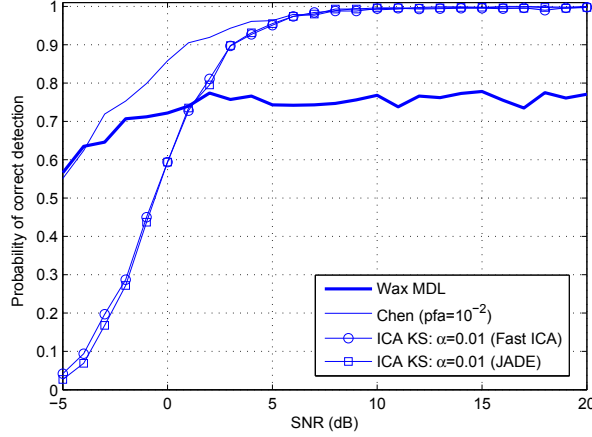
Fig. 3. Probability of correct detection of the number of sources($n_t = 2$) for a 2×2-antenna receiver and 1000 received samples per antenna.

Fig. 4. Probability of correct classification according to the number of transmitted signals n_t with a SNR equals to 10 dB. The receiver consists of 2×2 antennas and intercepts 1000 samples.

in presence of correlated noise, the Wax algorithm totally fails even for high SNR. The performances of the method proposed by Chen are better than those of our method at low SNR but are quite similar for a SNR higher than 5dB.

The figure 4 exposes the performances of the different methods according to the number of communication signals for a SNR equals to 10dB and 2×2 receiver antennas. We can remark that the Chen algorithm is not able to detect more than 2 signals. For a fixed number of receive antennas, the main benefit of our algorithm against those proposed by Chen, is its ability to detect a larger number of communication signals. Nevertheless, the performances of our method decreases when n_t is close to n_r.

5. CONCLUSION

In order to detect the number of communication signals from a sensor array, we proposed a new approach based on the extraction of non-gaussian components from the received samples. The extraction of non-gaussian components is realized by an ICA algorithm. Then, the number of communication signals corresponds to the number of extracted components which are non-gaussian circular. This number is estimated with Kolmogorov Smirnov tests. Simulation results proved the robustness of our proposed method even under condition of correlated noise.

6. REFERENCES

[1] A. James, *Test of equality of the latent Roots of the covariance and correlation matrices*, Biometrica, Vol. 43, p 128-136, 1956.
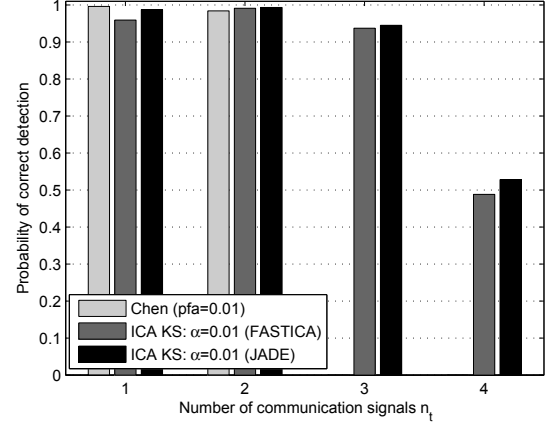
[2] M. Wax, and T. Kailath, *Detection of Signals by Information Theoretic Criteria*, IEEE Transactions on Acoustics, Speech, and Signal processing, Vol. ASSP-33, NO. 2, 1985.

[3] W. Chen, J.P. Reilly, and K.M. Wong, *Detection of the number of signals in noise with banded covariance matrices*, IEEE Radar, Sonar and Navigation, p 289-294, 1996.

[4] H. Sawada, R. Mukai, S. Araki, and S. Makino, *Estimating the number of sources using independent component analysis*, Acoust. Sci. & Tech Letter, Vol. 5, p 450-452, 2005.

[5] P. Hirschauer, J. Grouffaud, P. Larzabal, and H. Clergeot, *A neural network for model order selection in signal processing*, IEEE Neural Networks, Vol. 6, p 3057-3061, 1995.

[6] A. Hyvarinen, J. Karhunen, and E. Oja, *Independant Component Analysis*, Wiley & Sons, 2001.

[7] J-F. Cardoso, and A. Souloumiac, *Blind Beamforming for non gaussian signals*, IEE T, Vol. 140, No 6 362-370, 1993.

[8] E. Bingham, and A. Hyvarinen, *A fast fixed point algorithm for independent component analysis of complex valued signals*, Int Journal on Neural Systems, Vol. 10, No 1 27-47, 1999.

[9] P. Chevalier, L. Albera, P. Comon, and A. Ferreol *Comparative performance analysis of eight blind source separation method on radiocommunications signals*, IEEE International Joint Conference on Neural Networks, Hungary, 2004.